



OIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/037,677

DATE: 02/07/2002

TIME: 19:15:31

Input Set : N:\Crf3\RULE60\10037677.txt
Output Set: N:\CRF3\02072002\J037677.raw

2 <110> APPLICANT: Schellenberger, Volker
3 Liu, Amy D.
4 Selifonova, Olga V.
6 <120> TITLE OF INVENTION: Directed Evolution of Microorganisms
8 <130> FILE REFERENCE: GC560
10 <140> CURRENT APPLICATION NUMBER: 10/037,677
11 <141> CURRENT FILING DATE: 2001-10-23
13 <150> PRIOR APPLICATION NUMBER: 09/314,847
14 <151> PRIOR FILING DATE: 1999-05-19
16 <160> NUMBER OF SEQ ID NOS: 15
18 <170> SOFTWARE: FastSEQ for Windows Version 3.0
20 <210> SEQ ID NO: 1
21 <211> LENGTH: 741
22 <212> TYPE: DNA
23 <213> ORGANISM: E. coli
25 <400> SEQUENCE: 1
26 atgaccgcta tgagcactgc aattacacgc cagatcggtc tcgataccga aaccaccgg 60
27 atgaaccaga ttgggtcgca ctatgaaggc cacaagatca ttgagatgg tgcgttgaa 120
28 gtggtaacc gtcgcctgac gggcaataac ttccatgttt atctcaaacc cgatcggtg 180
29 gtggatccgg aaggccctgg cgtacatggt attgccgatg aatttttgcg cgataagccc 240
30 acgtttgccg aagtagccga tgagttcatg gactatattc gccgcgcgg 60
31 cataacgcag cggtcgatat cggctttatg gactacgatg ttgcgttgct taagcgcat 360
32 attccgaaga ccaataactt ctgtaaaggc accgatagcc ttgcgttggc gagaaaaatg 420
33 ttcccggta agcgcacacg cctcgatgatg ttatgtgatc gctacgaaat agataacatg 480
34 aaacgaacgc tgcacggggc attactcgat gcccagatcc ttgcggaaat ttatctggcg 540
35 atgaccgggt gtcaaacgtc gatggcttt gcgttggaaag gagagacaca acagcaacaa 600
36 ggtgaagcaa caattcagcg cattgtacgt caggcaagta agttacgcgt tgtttttgcg 660
37 acagatgaag agattgcacgc tcatgaagcc cgtctcgatc tggtgcagaa gaaaggcgg 720
38 agttgcctct ggcgagcata a 741
40 <210> SEQ ID NO: 2
41 <211> LENGTH: 246
42 <212> TYPE: PRT
43 <213> ORGANISM: E. coli
45 <400> SEQUENCE: 2
46 Met Thr Ala Met Ser Thr Ala Ile Thr Arg Gln Ile Val Leu Asp Thr
47 1 5 10 15
48 Glu Thr Thr Gly Met Asn Gln Ile Gly Ala His Tyr Glu Gly His Lys
49 20 25 30
50 Ile Ile Glu Ile Gly Ala Val Glu Val Val Asn Arg Arg Leu Thr Gly
51 35 40 45
52 Asn Asn Phe His Val Tyr Leu Lys Pro Asp Arg Leu Val Asp Pro Glu
53 50 55 60
54 Ala Phe Gly Val His Gly Ile Ala Asp Glu Phe Leu Leu Asp Lys Pro

ENTERED

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/037,677

DATE: 02/07/2002
TIME: 19:15:31

Input Set : N:\Crf3\RULE60\10037677.txt
Output Set: N:\CRF3\02072002\J037677.raw

55	65	70	75	80
56	Thr Phe Ala Glu Val Ala Asp Glu Phe Met Asp Tyr Ile Arg Gly Ala			
57	85	90	95	
58	Glu Leu Val Ile His Asn Ala Ala Phe Asp Ile Gly Phe Met Asp Tyr			
59	100	105	110	
60	Glu Phe Ser Leu Leu Lys Arg Asp Ile Pro Lys Thr Asn Thr Phe Cys			
61	115	120	125	
62	Lys Val Thr Asp Ser Leu Ala Val Ala Arg Lys Met Phe Pro Gly Lys			
63	130	135	140	
64	Arg Asn Ser Leu Asp Ala Leu Cys Ala Arg Tyr Glu Ile Asp Asn Ser			
65	145	150	155	160
66	Lys Arg Thr Leu His Gly Ala Leu Leu Asp Ala Gln Ile Leu Ala Glu			
67	165	170	175	
68	Val Tyr Leu Ala Met Thr Gly Gly Gln Thr Ser Met Ala Phe Ala Met			
69	180	185	190	
70	Glu Gly Glu Thr Gln Gln Gln Gly Glu Ala Thr Ile Gln Arg Ile			
71	195	200	205	
72	Val Arg Gln Ala Ser Lys Leu Arg Val Val Phe Ala Thr Asp Glu Glu			
73	210	215	220	
74	Ile Ala Ala His Glu Ala Arg Leu Asp Leu Val Gln Lys Lys Gly Gly			
75	225	230	235	240
76	Ser Cys Leu Trp Arg Ala			
77	245			
79	<210> SEQ ID NO: 3			
80	<211> LENGTH: 1164			
81	<212> TYPE: DNA			
82	<213> ORGANISM: E. blattae			
84	<400> SEQUENCE: 3			
85	atgagctatc gtatgttga ttatctgggt ccaaatgtga acttctttgg cccggcgcc	60		
86	gtttctgttg ttggccagcg ctgcgcagctg ctggggggta aaaaagccct gctggtgacc	120		
87	gataagggcc tgcgcgcccc taaagacggt gctgtcgatc agaccgtgaa gcacctgaaa	180		
88	gccgcggta ttgaggtgtt cattttcgac ggggtcgagc cgaacccgaa agacaccaac	240		
89	gtgctcgacg gcctggccat gttccgtaaa gagcagtgcg acatgataat caccgtcgcc	300		
90	ggcggcagcc cgcacgactg cgttaaaggc attggatttgcg cggccaccca cccgggtgat	360		
91	ctgtacagct atgcccgtat cggaaacactc accaaccgc tgccgccccat tattgggtc	420		
92	aacaccaccc cgccggaccgc cagcgaagtc acccggcaact gctgtcgatc taacaccaaa	480		
93	accaaaagtaa aatttgtat tgcagctgg cgcaacctgc cttccgtctc cattaacgat	540		
94	ccgctgctga tgcgtccaa gcccgcggg ctgaccgcg ccacccgtat ggatgccctg	600		
95	accacacgcgg tagaggccctatctccaaa gacgccaacc cgggttaccgc tgccctctgct	660		
96	attcaggcca tcaaactgtat tgcaccaccaac ttgcgcagg cctgtcgccct ggggaccaac	720		
97	ctcaaagccc gtgaaaacat ggcctgcgc tctctgtgg cccggatggc ctttaacaac	780		
98	gcacacctgg gctatgttca cgcacatggct caccagctgg ggggcctgtat cgacatggcc	840		
99	cacgggtgg cgaacgcggc cctgctgccc catgtctgcg gctataaccc gattgccaac	900		
100	ccggaaaaat ttgcgcgatcatcgccacccat atggggaaaa acaccacccgg tcttccacc	960		
101	atggacgcag cggagctggc catcagcgcc attgcccgtc tgcgtaaaga tgcgtggatc	1020		
102	ccgcagcacc tgcgtgaact ggggtaaaaa gagggccact tcccgatcat ggcagaaatg	1080		
103	gcacccgtaaag acggcaacgc cttctctaaac cgcaccaag ggaacgaaaa agagattgcc	1140		
104	gacatttcc gccaggcatt ctga	1164		
106	<210> SEQ ID NO: 4			

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/037,677

DATE: 02/07/2002
TIME: 19:15:31

Input Set : N:\Crf3\RULE60\10037677.txt
Output Set: N:\CRF3\02072002\J037677.raw

107 <211> LENGTH: 387
108 <212> TYPE: PRT
109 <213> ORGANISM: E. blatte
111 <400> SEQUENCE: 4
112 Met Ser Tyr Arg Met Phe Asp Tyr Leu Val Pro Asn Val Asn Phe Phe
113 1 5 10 15
114 Gly Pro Gly Ala Val Ser Val Val Gly Gln Arg Cys Gln Leu Leu Gly
115 20 25 30
116 Gly Lys Lys Ala Leu Leu Val Thr Asp Lys Gly Leu Arg Ala Ile Lys
117 35 40 45
118 Asp Gly Ala Val Asp Gln Thr Val Lys His Leu Lys Ala Ala Gly Ile
119 50 55 60
120 Glu Val Val Ile Phe Asp Gly Val Glu Pro Asn Pro Lys Asp Thr Asn
121 65 70 75 80
122 Val Leu Asp Gly Leu Ala Met Phe Arg Lys Glu Gln Cys Asp Met Ile
123 85 90 95
124 Ile Thr Val Gly Gly Ser Pro His Asp Cys Gly Lys Gly Ile Gly
125 100 105 110
126 Ile Ala Ala Thr His Pro Gly Asp Leu Tyr Ser Tyr Ala Gly Ile Glu
127 115 120 125
128 Thr Leu Thr Asn Pro Leu Pro Pro Ile Ile Ala Val Asn Thr Thr Ala
129 130 135 140
130 Gly Thr Ala Ser Glu Val Thr Arg His Cys Val Leu Thr Asn Thr Lys
131 145 150 155 160
132 Thr Lys Val Lys Phe Val Ile Val Ser Trp Arg Asn Leu Pro Ser Val
133 165 170 175
134 Ser Ile Asn Asp Pro Leu Leu Met Ile Gly Lys Pro Ala Gly Leu Thr
135 180 185 190
136 Ala Ala Thr Gly Met Asp Ala Leu Thr His Ala Val Glu Ala Tyr Ile
137 195 200 205
138 Ser Lys Asp Ala Asn Pro Val Thr Asp Ala Ser Ala Ile Gln Ala Ile
139 210 215 220
140 Lys Leu Ile Ala Thr Asn Leu Arg Gln Ala Val Ala Leu Gly Thr Asn
141 225 230 235 240
142 Leu Lys Ala Arg Glu Asn Met Ala Cys Ala Ser Leu Leu Ala Gly Met
143 245 250 255
144 Ala Phe Asn Asn Ala Asn Leu Gly Tyr Val His Ala Met Ala His Gln
145 260 265 270
146 Leu Gly Gly Leu Tyr Asp Met Ala His Gly Val Ala Asn Ala Val Leu
147 275 280 285
148 Leu Pro His Val Cys Arg Tyr Asn Leu Ile Ala Asn Pro Glu Lys Phe
149 290 295 300
150 Ala Asp Ile Ala Thr Phe Met Gly Glu Asn Thr Thr Gly Leu Ser Thr
151 305 310 315 320
152 Met Asp Ala Ala Glu Leu Ala Ile Ser Ala Ile Ala Arg Leu Ser Lys
153 325 330 335
154 Asp Val Gly Ile Pro Gln His Leu Arg Glu Leu Gly Val Lys Glu Ala
155 340 345 350
156 Asp Phe Pro Tyr Met Ala Glu Met Ala Leu Lys Asp Gly Asn Ala Phe

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/037,677

DATE: 02/07/2002
TIME: 19:15:31

Input Set : N:\Crf3\RULE60\10037677.txt
Output Set: N:\CRF3\02072002\J037677.raw

157 355 360 365
158 Ser Asn Pro Arg Lys Gly Asn Glu Lys Glu Ile Ala Asp Ile Phe Arg
159 370 375 380
160 Gln Ala Phe
161 385
163 <210> SEQ ID NO: 5
164 <211> LENGTH: 12
165 <212> TYPE: DNA
166 <213> ORGANISM: Artificial Sequence
168 <220> FEATURE:
169 <223> OTHER INFORMATION: wild type mutD gene
171 <400> SEQUENCE: 5
172 atgaccgcta tg 12
174 <210> SEQ ID NO: 6
175 <211> LENGTH: 11
176 <212> TYPE: DNA
177 <213> ORGANISM: Artificial Sequence
179 <220> FEATURE:
180 <223> OTHER INFORMATION: pOS100 mutD mutated gene
182 <400> SEQUENCE: 6
183 ttgacgctt g 11
185 <210> SEQ ID NO: 7
186 <211> LENGTH: 12
187 <212> TYPE: DNA
188 <213> ORGANISM: Artificial Sequence
190 <220> FEATURE:
191 <223> OTHER INFORMATION: pOS101 mutD mutated gene
193 <400> SEQUENCE: 7
194 gtgaccgctg tg 12
196 <210> SEQ ID NO: 8
197 <211> LENGTH: 11
198 <212> TYPE: DNA
199 <213> ORGANISM: Artificial Sequence
201 <220> FEATURE:
202 <223> OTHER INFORMATION: pOS102 mutD mutated gene
204 <400> SEQUENCE: 8
205 gtgccgctgt g 11
207 <210> SEQ ID NO: 9
208 <211> LENGTH: 12
209 <212> TYPE: DNA
210 <213> ORGANISM: Artificial Sequence
212 <220> FEATURE:
213 <223> OTHER INFORMATION: pOS104 mutD mutated gene
215 <400> SEQUENCE: 9
216 ttgaccgctt tg 12
218 <210> SEQ ID NO: 10
219 <211> LENGTH: 55
220 <212> TYPE: DNA
221 <213> ORGANISM: Artificial Sequence

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/037,677

DATE: 02/07/2002
TIME: 19:15:31

Input Set : N:\Crf3\RULE60\10037677.txt
Output Set: N:\CRF3\02072002\J037677.raw

223 <220> FEATURE:
224 <223> OTHER INFORMATION: pOS105 mutD mutated gene
226 <400> SEQUENCE: 10
227 gtgaccgctg tgagcacttg caattacacg ccagatcggtt ctcgataccg aaatc 55
229 <210> SEQ ID NO: 11
230 <211> LENGTH: 11
231 <212> TYPE: DNA
232 <213> ORGANISM: Artificial Sequence
234 <220> FEATURE:
235 <223> OTHER INFORMATION: pOS106 mutD mutated gene
237 <400> SEQUENCE: 11
238 gtgaccgctt g 11
240 <210> SEQ ID NO: 12
241 <211> LENGTH: 28
242 <212> TYPE: DNA
243 <213> ORGANISM: Artificial Sequence
245 <220> FEATURE:
246 <223> OTHER INFORMATION: primer
248 <400> SEQUENCE: 12
249 cgcctccagc gcgacaatag cggccatc 28
251 <210> SEQ ID NO: 13
252 <211> LENGTH: 27
253 <212> TYPE: DNA
254 <213> ORGANISM: Artificial Sequence
256 <220> FEATURE:
257 <223> OTHER INFORMATION: primer
259 <400> SEQUENCE: 13
260 ccgactgaac taccgctccg cgttgtg 27
262 <210> SEQ ID NO: 14
263 <211> LENGTH: 36
264 <212> TYPE: DNA
265 <213> ORGANISM: Artificial Sequence
267 <220> FEATURE:
268 <223> OTHER INFORMATION: primer
270 <400> SEQUENCE: 14
271 tctgatacgg gatcctcaga atgcctggcg gaaaat 36
273 <210> SEQ ID NO: 15
274 <211> LENGTH: 42
275 <212> TYPE: DNA
276 <213> ORGANISM: Artificial Sequence
278 <220> FEATURE:
279 <223> OTHER INFORMATION: primer
281 <400> SEQUENCE: 15
282 gcgccgtcta gaattatgag ctatcgatg tttgattatc tg 42

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/037,677

DATE: 02/07/2002

TIME: 19:15:32

Input Set : N:\Crf3\RULE60\10037677.txt
Output Set: N:\CRF3\02072002\J037677.raw